Jour. Fac. Sci., Hokkaido Univ., Ser. VII (Geophysics), Vol. 11, No. 4, 725-738, 2000.

Modern Signal Extraction Methods in Computational Seismology

Tetsuo Takanami

The Institute of Seismology and Volcanology, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan

and

Genshiro Kitagawa

The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan

(Received November 30, 1999)

Abstract

The record of Earth motion obtained by a seismograph contains information about the nature of seismic source that generated the motion as well as properties of the media through which the seismic disturbances were propagated. Our problem is to extract such information from the observed record. However the seismograph is under continuous influence of a variety of natural forces such as the effect of microtremor by the natural sources of winds and waves in the sea and by the common artificial sources by human activity. Since it is almost impossible to describe the response to these noise inputs precisely, for automatic processing of seismic data, proper statistical modeling is necessary. In this paper, we show the several specific examples of time series modeling for signal extraction problems related to seismology. Namely, we consider the extraction of small seismic signal from noisy data by the local likelihood method and the self-organizing state space modeling. The extraction of weak signals from noisy data has been shown to exemplify the power of the above procedures.

1. Introduction

Earthquake networks are designed to record as many earthquakes as nature and instruments permit, and they can produce an immense volume of data. For small earthquake network in areas of high seismicity, or for large microearthquake networks, processing voluminous data can be a serious problem. Especially, since 1979, several areas in Japan have been specified as priority observation areas and well-equipped nation-wide seismological net-

work (Hamaguchi and Suzuki, 1979) was established. However, the larger volume of microearthquake data requires application of modern data processing techniques. The data collected by seismologists are quite diversified because of the variety of seismic sources; the great range in source size, travel distance, and receiver-array dimension; and the different types of measurements made with seismic waves. In practice, moreover, the seismic signals observed by the earthquake networks are contaminated by various kinds of natural sources, such as microtremors, microseisms, waves in the sea and a variety of human induced sources. Particularly, for the processing of increasingly many microearthquakes, it thus becomes necessary to develop a computationally efficient statistical method that can extract seismic signal from noisy data. For this purpose, a more sophisticated procedure which can handle very noisy data is required. Recently, we also discussed the signal extraction problems related to seismology (Kitagawa et. al., 1999). Namely, they considered 1) the estimation of the arrival time of seismic signal, 2) the extraction of small seismic signal from noisy data, 3) the detection of the coseismic effect in groundwater level data contaminated by various effects from air pressure etc., and the estimation of changing spectral characteristic of seismic record. The models can be expressed in state space model form, which enable to apply computationally efficient Kalman filter or extension of it.

In this paper, we reconsider the extraction of small seismic signal from noisy data. We apply the nonstationary structural time series models to the several seismic signals observed by the seismological network, namely three feeble seismic signals of fore-shock and after-shock of the 1982 Urakawa-Oki earthquake and small later phase of SxS after S wave recorded on the high dense seismological network in the Nikko area. The information criterion AIC (Akaike, 1973) is used for the evaluation of the goodness of the model which facilitates to establish automatic signal extraction procedures.

2. Extraction methods of seismic signal from noisy data

The problem of extracting a signal of microearthquake from noisy data is considered here. As noted before, the cumulative number of earthquakes increases exponentially as the magnitude decreases. Therefore, by developing a method for analyzing small signal with smaller amplitudes, we can get increasingly more information about the seismicity of the region. However, the Earth's surface is actually under continuous disturbances due to a variety of natural forces and human induced sources. Therefore, if the amplitude of the

earthquake signal is very small, it will be quite difficult to distinguish it from the background noise. In this section, we consider a method of extracting small seismic signals in the presence of relatively large background noise.

2.1 The time series models for the extraction of seismic signal

In this section we assume that observed seismogram is consisted of three additive components, the background noise component, the seismic signal and the observation noise. Namely, we assume that the observed seismogram y_n is expressed as

$$y_n = r_n + s_n + \varepsilon_n, \tag{1}$$

where r_n, s_n and ε_n represent background noise, the signal and the observation noise, respectively.

The background noise component r_n is generated by various noise sources and sometimes has significant power spectrum in low frequency ranges. Therefore we approximate the process by an autoregressive model

$$r_n = \sum_{i=1}^{m} a_i r_{n-1} + u_n, \qquad u_n \sim N(0, \tau_1^2), \tag{2}$$

where the AR order m, the AR coefficients a_i and the innovation variance τ_1^2 are unknown parameters. On the other hand, the seismic signal component s_n is consisted mainly of the P-wave and S-wave and will have several spectral peaks. Therefore, we shall also approximate the seismic signal component by another autoregressive model

$$s_n = \sum_{i=1}^{l} b_i s_{n-i} + v_n, \qquad v_n \sim N(0, \tau_2^2), \tag{3}$$

where the AR order ℓ, and the AR coefficients b_i and the innovation variance τ_2^2 are unknown parameters. For the moments, we treat τ_2^2 as an unknown constant, it is actually time-varying. We assume that the observation noise component ε_n is an independent sequence and is a Gaussian white noise process $\varepsilon_n \sim N(0, \sigma^2)$ (Kitagawa and Takanami, 1985).

The models in (1), (2) and (3) can be combined in the state space model form

$$\begin{aligned} x_n &= F x_{n-1} + G w_n, \\ y_n &= H x_n + \varepsilon_n \end{aligned} \tag{4}$$

where $(m+\ell)$-dimensional state vector x_n is defined by $x_n = (r_n, \ldots, r_{n-m+1}, s_n, \ldots, s_{n-\ell+1})^T$, and $w_n = (u_n, v_n)^T$ is a two dimensional system noise. F, G and H are respectively $(m+\ell)\times(m+\ell)$ matrix and $(m+\ell)\times 2$ and $1+(m+\ell)$ vectors defined by

$$F=\left[\begin{array}{cccc|cccc} a_1 & a_2 & \cdots & a_m & & & & \\ 1 & & & & & & & \\ & \ddots & & & & 0 & & \\ & & 1 & & & & & \\ \hline & & & & b_1 & b_2 & \cdots & b_\ell \\ & & & & 1 & & & \\ & 0 & & & & \ddots & & \\ & & & & & & 1 & \end{array}\right], \quad G=\left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ \hline 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{array}\right]$$

$$H=[\,1\ 0\cdots 0\ |\ 1\ 0\cdots 0\,]. \tag{5}$$

The variance of w_n is given by $Q_n=\mathrm{Diag}\ \{\tau_1^2, \tau_{2n}^2\}$.

Since the state vector x_n contains both the background noise component r_n and the seismic component s_n, this state space model shows that if we can estimate the state x_n from the observed data, it immediately provides us with the estimates of background noise and the seismic signal.

2.2 *Extraction of the Signal by the Kalman Filter and Smoother*

As mentioned in the previous subsections, the AR orders, m and ℓ, the AR coefficients a_j and b_j and the variances σ^2, τ_1^2 and τ_2^2 are actually unknown parameters that need to be estimated by the observations. However, for the moment we assume that they are given and consider the estimation of the state vector x_n. Note that since the state vector is $(m+\ell)$-dimensional, we have at least $(m+\ell)\times N$ unknowns for N observations and thus it is almost impossible to obtain by the ordinary least squares method.

Denote the estimate of the state given the observations $y_1, \ldots, y_t$ and its variance-covariance matrix, by $x_{n|t}$ and $V_{n|t}$, respectively. Then the one-step-ahead predictor $x_{n|n-1}$ and the filter $n_{n|n}$ can be obtained recursively by the following Kalman filter (Anderson and Moor, 1979):

Prediction

$$\begin{aligned} x_{n|n-1}&=Fx_{n-1|n-1} \\ V_{n|n-1}&=FV_{n-1|n-1}F^T+GQ_nG^T \end{aligned} \tag{6}$$

Filter

$$\begin{aligned} K_n&=V_{n|n-1}H^T(HV_{n|n-1}H^T+R)^{-1} \\ x_{n|n}&=x_{n|n-1}+K_n(y_n-Hx_{n|n-1}) \\ V_{n|n}&=(I-K_nH)\,V_{n-1|n-1}. \end{aligned} \tag{7}$$

Here the initial values $x_{0|0}$ and $V_{0|0}$ can be determined from the assumed model as follows.

If the AR model is given, it can be easily seen that the unconditional mean value of the state vector is zero. Therefore, we put $x_{0|0}=0$. On the other hand, the autocovariance function of the AR process is obtained as the solution to the Yule-Walker equation

$$C_0=\sum_{j=1}^{m} a_j C_j+\sigma^2$$
$$C_k=\sum_{j=1}^{m} a_j C_{j-k} \qquad k=1, \ldots, m \tag{8}$$

Then the initial variance covariance matrix of $(r_n, \ldots, r_{n-m+1})$ is given by

$$V_{0|0}=\begin{bmatrix} C_0 & C_1 & \cdots & C_{m-1} \\ C_1 & C_0 & \cdots & C_{m-2} \\ \vdots & \vdots & \vdots & \vdots \\ C_{m-1} & C_{m-2} & \cdots & C_0 \end{bmatrix} \tag{9}$$

Then the above step of the Kalman filter can be repeated to obtain the one-step-ahead prediction and the filter can be repeated as long as the observations are obtained.

The final estimates of the state vectors $x_{n|N}$, which is the estimates of the state based on the entire data, are obtained by the following fixed interval smoothing algorithm (Anderson and Moore, 1979);

$$A_n=V_{n|n}F^T V_{n+1|n}^{-1}$$
$$x_{n|N}=x_{n|n}+A_n(x_{n+1|N}-x_{n+1|n}) \tag{10}$$
$$V_{n|N}=V_{n|n}+A_n(V_{n+1|N}-V_{n+1|n})A_n^T$$

Then the first component of $x_{n|N}$ is the estimate of the background noise component $r_{n|N}$ and the $m+1$-st component of $x_{n|N}$ is the estimate of seismic signal $s_{n|N}$.

2.3 Estimation of the Model Parameters

In actual estimation, the parameters of the model θ is unknown and need to be estimated from the data. Under the assumption of stationarity of the background noise, the parameters of the model a_i and τ_1^2 can be estimated by fitting the AR model with observation noise to a part of the data where the seismic signal apparently does not exist;

$$y_n=r_n+\varepsilon_n$$
$$r_n=\sum_{i=1}^{m} a_i r_{n-i}+u_n. \tag{11}$$

The state space representation for this model can be obtained by considering the

special case when $l=0$ in (3). The log-likelihood of this AR(m) plus noise model is obtained by

$$l(\theta_m)=-\frac{N}{2}\log 2\pi-\frac{1}{2}\sum_{n=1}^{N}\log r_n-\frac{1}{2}\Sigma\frac{\varepsilon_n^2}{r_n}, \tag{12}$$

where $\varepsilon_n=y_n-Hx_{n|n-1}$ and $r_n=Hv_{n|n-1}H^T+\sigma^2$ (Jones (1980)).

Then, by fixing the obtained parameters for the background noise model, the parameters of the seismic signal model b_i and τ_2^2 are obtained by fitting the model (1), (2) and (3). This two-step estimation procedure will be reasonable since the background noise can be considered stationary for a certain time interval.

2.4 Estimation of the Time Varying Variance

The variance of the autoregressive model for the signal corresponds to the amplitude of the seismic signal and is actually time varying. Namely, the variance, τ_2^2, is most zero before the seismic signal arrives, and when the seismic signal arrives it becomes large depending on the amplitude of the signal and then goes back to zero as the tail of the seismic signal dies out. This variance parameter plays the role of a signal to noise ratio, and the estimation of this parameter is the key to the success of extracting the seismic signal.

In earlier paper (Kitagawa and Takanami, 1985), this was achieved by defining the local likelihood which evaluates the goodness of predetermined candidates of variance and finding the best one for each time instance.

Recently, a self-organizing state space model is proposed for simultaneous estimation of the state and the parameter of the model (Kitagawa, 1998). In this method, the original state vector x_n is augmented with the time-varying parameter θ_n as

$$z_n=\begin{bmatrix} x_n \\ \theta_n \end{bmatrix} \tag{13}$$

For our problem where only the variance of the seismic signal component τ_2^2 is time-varying, θ_n is defined by

$$\theta_n=\log_{10}\tau_{2n}^2. \tag{14}$$

The log transformation for the variance is used to assure the positively of τ_{2n}^2. We further assume that this parameter θ_n changes according to the random walk model

$$\log_{10}\tau_{2,n}^2=\log_{10}\tau_{2,n-1}^2+\eta_n, \tag{15}$$

where η_n is the Gaussian white noise with $\eta_n \sim N(0, \xi^2)$. Note that it is possible to use other model such as the second order trend model or an autoregressive model.

The state space model for this augmented state is easily obtained from the original state space model for x_n and (15);

$$\begin{bmatrix} x_n \\ \log_{10}\tau_{2,n}^2 \end{bmatrix} = \begin{bmatrix} F & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_{n-1} \\ \log_{10}\tau_{2,n-1}^2 \end{bmatrix} + \begin{bmatrix} G & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_n \\ \eta_n \end{bmatrix}. \tag{16}$$

Then by applying the nonlinear non-Gaussian smoother based on the Monte Carlo method (Kitagawa (1996), we can estimate the augmented state. Since the augmented state contains x_n and θ_n, this means that the marginal posterior density of x_n and θ_n can be obtained simultaneously.

3. Examples: Separating weak seismic signal from noisy data

Using the AR model for the background noise, the AR model was fitted to the data where the seismic signal apparently exists. Figure 1A shows the record of U-D component of a fore-shock of the 1982 Urakawa earthquake recorded at station Hidaka of the seismological network of Hokkaido University (Kitagawa and Takanami, 1985; Takanami, 1991). We apply the local likelihood method for the estimation of changing variance (Kitagawa and Takanami, 1985). The background noise was estimated using 400 observations before the earthquake. Figure 1B, 1C, 1D and 1E respectively illustrate the extracted background noise, the seismic signal, the observation noise, and the time varying variance. In spite of the smaller signal, we can get quite good decomposition. The existence of earthquake signal that is not so clear in the original recorded is highlighted. P- and S-arrives can be clearly identified around at $n=400$ and at $n=680$, respectively. By this decomposition, it becomes clear that the background noise and observation noise are stationary over time, respectively.

Figure 2A shows the records of the E-W component of a local earthquake recorded by the high dense seismological network in the Nikko area (Urabe et al., 1994: Takanami et al., 1995). The small later phase which reflected from the strong reflector that exists beneath the Nikko area is assigned to SxS (e.g., Iidaka, 1994; Matsumoto, 1994; Uehira, 1994). We also apply the local likelihood method for the estimation of changing variance. The background noise model was estimated using 600 observations before the SxS. Figure 2B shows a magnified plot of original record of SxS contaminated with scattered coda

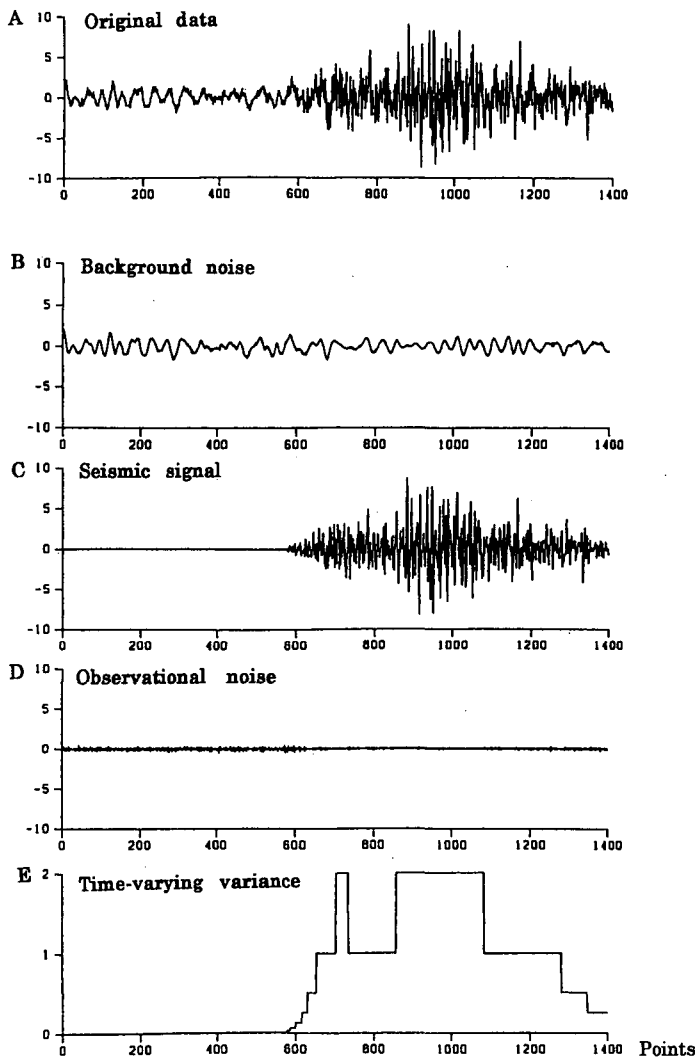

Fig. 1. Decomposition by local likelihood. A: original wave-form of the foreshock of 1982 Urakawa-Oki earthquake (Station=Hidaka, component=NS, origin time=9:33, March 21, 1982, depth=31.0 km, M=2.3), B: background noise, C: seismic signal, D: observational noise, E: the variance of the seismic signal model. This figure is compounded from the figures of 1 to 3 in the paper by Kitagawa and Takanami (1985) and the seismogram is filed in ISM data 43-3-01 (Takanami, 1991).

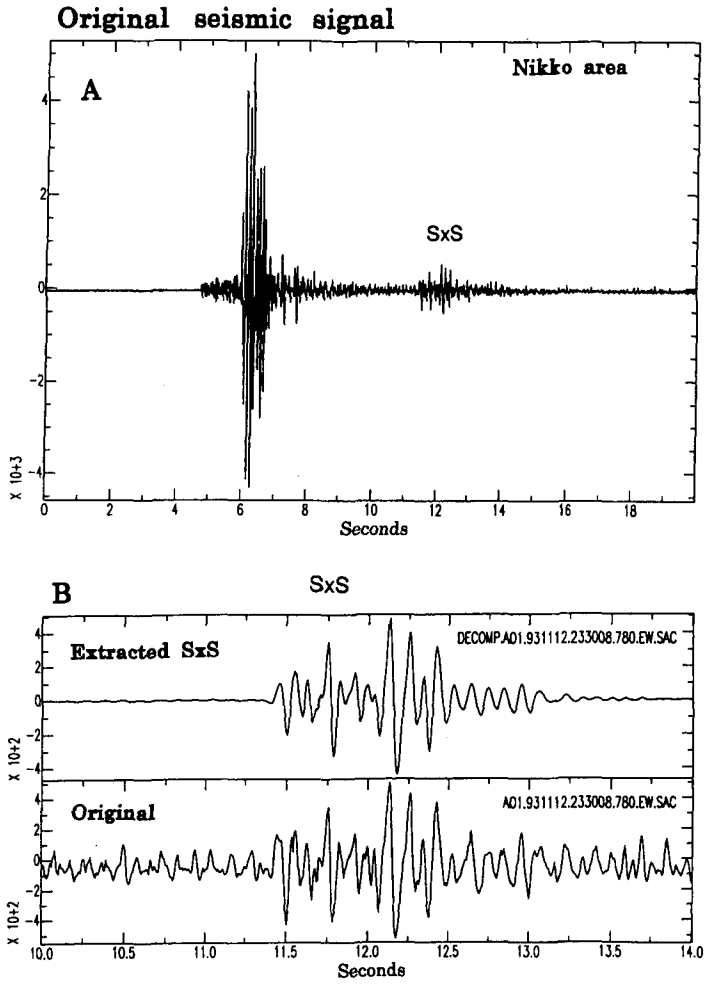

Fig. 2. Extraction of SxS from the S-coda waves by local likelihood. A : original wave-form of shallow earthquake in Nikko area (station = A01, component = EW, origin time = 23 : 30, November 12, 1993, depth = 3.5 km, M = 0.42), B : Original noisy later phase around SxS and extracted SxS signal.

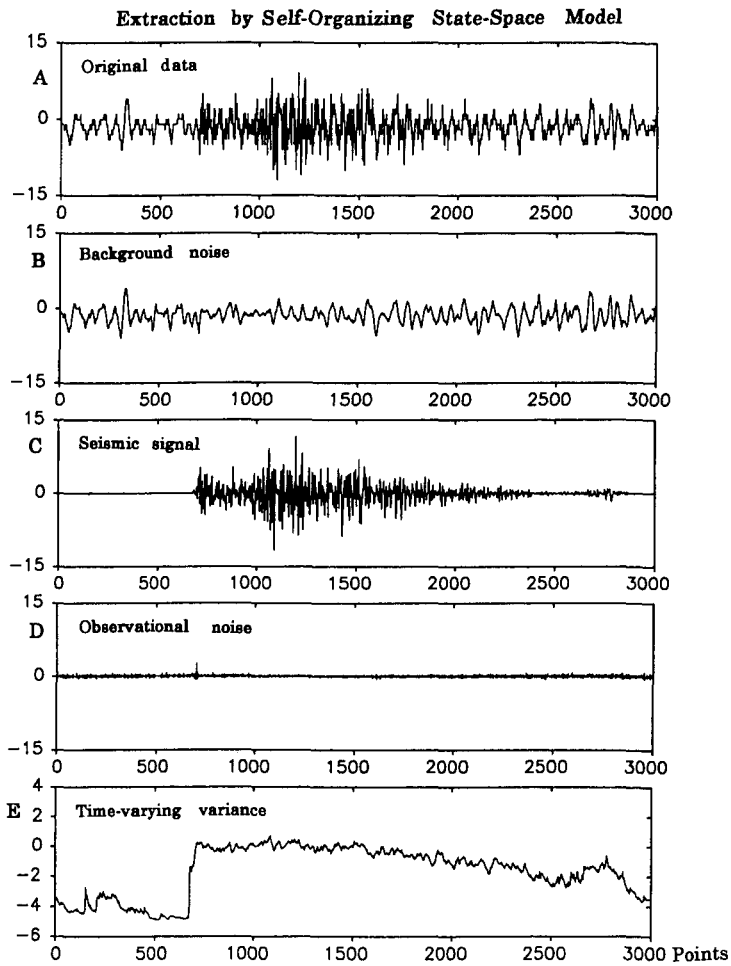

Fig. 3. Decomposition by self-organizing state-space model. A: original waveform of the fore-shock of 1982 Urakawa-Oki earthquake (Station=Erimo, component=NS, origin time=07:45, March 21, 1982, depth=31.0 km, M=1.9), B: background noise, C: seismic signal, D: observational noise, E: the variance of the seismic signal model. This figure is identical with figure of 3 in the paper by Kitagawa *et al.* (1999).

waves (lower) and the extracted SxS by the local likelihood method (upper). SxS wave is decomposed clearly with this method.

Next, we consider the self-organizing method for the general nonlinear non-Gaussian state-space model proposed by Kitagawa (1998). Here we apply a self-organizing state space model for the estimation of the time-varying variance by the Monte Carlo smoothing. The variance of the system noise was chosen by maximizing the log-likelihood on a coarse grid (see Fig. 4 in Kitagawa, 1998). Figure 3A shows the record of the N-S component of an aftershock of an earthquake (Takanami, 1991). The observed signal is minuscule relative to the background noise. We apply the self-organizing state space mode to the record. The background noise model was estimated using the first 600 observations. Figure 3B, 3C, and 3D respectively illustrate the extracted background noise, the seismic signal and the observation noise. It is clear that the background noise exists even after the arrival of the seismic signal. The background noise is almost homoscedastic in the whole interval, and also the seismic signal is clearly identified. The observation noise component is minuscule, looks homoscedastic, and is not shown here. Figure 3D shows the median of marginal posterior distribution of $\log_{10}\tau_{2n}^2$. The arrival of the seismic signal becomes apparent by this decomposition. Figure 3E shows the estimated variance function $\log_{10}\tau_{2n}^2$. It can be seen that the τ_{2n}^2 is approximately 10^{-4} for the background noise, increases to 10^0 by the arrival of P-wave and S-wave and then gradually decreases to the original noise level.

4. Discussion and Summary

When an earthquake signal arrives, the characteristics of the record of seismograms, such as the variances and the spectrum, change abruptly. For the estimation of the arrival time of a seismic signal, it is assumed that each of the seismogram before and after the arrival of the seismic wave is stationary and can be expressed by an autoregressive model (Takanami and Kitagawa, 1991 ; Takanami and Kitagawa, 1993 ; Takanami, 1999). By systematic use of statistical method, AIC criterion and computationally efficient estimation procedure, such as the least squares method based on Householder transformation, Kalman filter, it is possible to develop automatic procedures for the estimation of seismic signal.

In this paper, we extended the statistical time series model to the signal extraction problems related to seismology. We used the noisy data recorded by the microearthquake network. Namely, we considered the extraction of a weak

seismic signal from noisy data.

In this modeling the variance of the autoregressive model for the signal is related to the amplitude of seismic signal and is actually time varying. Namely, the variance is almost zero before the seismic signal arrives, becomes large depending on the amplitude of the signal and then goes back to zero as the tail of the seismic signal dies out. This variance parameter plays the role of a signal to noise ratio, and the estimation of this parameter is the key problem for the extraction of seismic signal.

In the earlier paper (Kitagawa and Takanami, 1985), it was realized by defining the local likelihood which evaluates the goodness of predetermined candidates of variance and find the best one for each time instance. Recently, a self-organizing state space model was successfully applied for the estimation of the time-varying variance (Kitagawa, 1998). We here showed three examples of time series modeling for signal extraction problem by the local likelihood (Kitagawa and Takanami, 1985) and the self-organizing state-space model (Kitagawa 1998). They were achieved by using the small three earthquakes occurred in the source area of the 1982 Urakawa-Oki earthquake and in the Nikko area where is an active volcanic region, respectively.

In the source area of the 1982 Urakawa-Oki earthquake, there occurred four small fore-shocks within two hours just before the main shock and their hypocenters are almost same as one of the main shock. In recently, such a fore-shock is commonly believed to be a quite important physical event which relates to the precursory slip before the coming main shock. However, it is very difficult to derive it such a small preseismic signal contaminated with strong noisy data.

On the other hand, in a small area in Nikko, Japan, about 200 seismometers were set and many micro-earthquakes were observed (Urabe et al., 1994). From this array data, it is expected to estimate precise underground structure, especially the shape and the size of the lump of magma. SxS waves found in this area are important phases for understanding the physical properties of the materials beneath the Nikko area. SxS wave generated from the reflector has an important information on the physical properties of the materials of reflectors distributed in the volcanic regions. But SxS is easily contaminated with the coda wave since it follows by the strong S wave. It is usually too small signal ratio to noise data (S/N ratio) to analyze for getting the information of the reflectors.

Examples such as adaptation to changes of the amplitude of a signal in seismic data is shown to exemplify the usefulness of the present extraction

methods.

References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. Second International Symposium on Information Theory, Academia Kiado, Budapest, 267-281.

Anderson, B.D.O, and J.B. Moore, 1979. *Optimal filtering*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 357 pp.

Hamaguchi, H. and Z. Suzuki, 1979. Automatic detection of onset time of microearthquake and its confidence. Report of Grant-in-aid for Scientific Research Project on Natural Disaster (No. 54-2), Ministry of Education, Science and Culture, 62-83 (in Japanese).

Iidaka, T., 1994. Reflectors found by the ultra high density network. In the Symposium on Heterogeneity and seismic activity in the crust organized by S. Horiuchi.

Jones, R.H., 1980. Maximum likelihood fitting of ARMA models to time series with missing observations. Technometrics, **22**, 389-395.

Kitagawa, G. and T. Takanami, 1985. Extraction of signal by a time series model and screening out micro earthquakes. Signal Processing, **8**, 303-314.

Kitagawa, G., 1996. Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. J. Compt. Graph. Statist, **5**, 1-25.

Kitagawa, G., 1998. A self-organizing state space model. J. Amer. Statist. Assoc., **93**, No. 443, 1203-1215.

Kitagawa, G., T. Takanami and N. Matsumoto, 1999. Signal extraction problems in seismology. Res. Memor. Inst. Inst. Statist. Math., **726**, 1-32.

Matsumoto, S., 1994. Distribution of reflectors in and near Nikko area. In the Symposium on Heterogeneity and seismic activity in the crust organized by S. Horiuchi.

Takanami, T., 1991. ISM data 43-3-01 : Seismograms of foreshocks of 1982 Urakawa-Oki earthquake. Ann. Inst. Statist. Math., **43**, No. 3, 605.

Takanami, T. and G. Kitagawa, 1991. Estimation of the arrival times of seismic waves by multivariate time series model. Ann. Inst. Statist. Math., **43**, No. 3, 407-433.

Takanami, T. and G. Kitagawa, 1993. Multivariate time-series model to estimate the arrival times of S-waves. Compt. Geosci., **19**, No. 2, 295-301.

Takanami T., S. Kodaira, M. Ichiyanagi, M. Okayama, H. Nagumo, M. Kasahara, T. Furumura and G. Kitagawa, 1995. Extraction of weak later-phases recorded on the high dense seismological network in the Nikko area. Abs. Japan Earth and Planetary Science Joint Meeting, 282 (in Japanese).

Takanami, T., 1999. High precision estimation of seismic wave arrival times. In the Practice Time Series Analysis, Akaike, H. and Kitagawa, G. eds., 79-94, Springer-Verlag, New York.

Uehira, K., 1994. Imaging of reflectors by the scattered waves. In the Symposium on Heterogeneity and seismic activity in the crust organized by S. Horiuchi.

Urabe and Group for the 1993 Joint Seismic Observation in and around the Nikko Area, 1994. A very dense seismic array observation for imaging the crust and seismic sources — The 1993 Joint seismic observation an and around Nikko area —. Abs. Japan Earth and Planetary Science Joint Meeting, 282 (in Japanese).